Published in final edited form as:

Spat Spatiotemporal Epidemiol. 2015 April ; 13: 31–40. doi:10.1016/j.sste.2015.04.004.

Do people really know what food retailers exist in their neighborhood? Examining GIS-based and perceived presence of retail food outlets in an eight-county region of South Carolina

Timothy L. Barnes^{a,b}, Bethany A. Bell^c, Darcy A. Freedman^d, Natalie Colabianchi^e, and Angela D. Liese^{a,*}

^aCenter for Research in Nutrition and Health Disparities and Department of Epidemiology and Biostatistics, Arnold School of Public Health, University of South Carolina, Columbia, SC, USA

^bObesity Prevention Center and Department of Epidemiology & Community Health, School of Public Health, University of Minnesota, Twin Cities, Minneapolis, MN, USA

^cCollege of Education, University of South Carolina, Columbia, SC, USA

^dDepartment of Epidemiology and Biostatistics, Case Western Reserve University, Cleveland, OH, USA

^eInstitute for Social Research, University of Michigan, Ann Arbor, MI, USA

Abstract

Measures of neighborhood food environments have been linked to diet and obesity. However, the appropriate measurement methods and how people actually perceive their food environments are still unclear. In a cross-sectional study of 939 adults, the perceived presence of food outlets was compared to the geographic-based presence of outlets within a participant's neighborhood, utilizing percent agreement and Kappa statistics. Perceived presence was based on survey-administered questions, and geographic-based presence was characterized using 1-, 2-, 3- and 5-mile (1-mile=1.6 km) Euclidean- and network-based buffers centered on each participant's residence. Analyses were also stratified by urban and non-urban designations. Overall, an individual's perceived neighborhood food environment was moderately correlated with the geographic-based presence of outlets. The performance of an individual's perception was most optimal using a 2- or 3-mile geographic-based neighborhood boundary and/or when the participant lived in a non-urban neighborhood. This study has implications for how researchers measure the food environment.

© 2015 Published by Elsevier Ltd.

*Corresponding Author: Angela D. Liese, PhD, MPH, Professor, Department of Epidemiology and Biostatistics, Center for Research in Nutrition and Health Disparities, Arnold School of Public Health, University of South Carolina, Discovery I, 461, 915 Greene Street, Columbia, SC, 29208, USA, liese@sc.edu, Phone: (803) 777-9414, Fax: (803) 777-2524.

Conflicts of Interest

The authors declare no conflicts of interest.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Keywords

Neighborhood; Food environment; Perceptions; GIS; Food outlets

1. Introduction

It has been suggested that neighborhood food environment, measured either objectively or subjectively, is associated with dietary intake (Caspi et al. 2012a). To date, geographic information systems (GIS) have been the most-utilized objective method to characterize neighborhood food environments (McKinnon et al., 2009; Charreire et al., 2010; Van et al., 2011; Thornton et al., 2011). However, it is unknown whether GIS-based measures are the most appropriate means of defining an individual's food environment (Caspi et al., 2012b, 2012a; Mujahid et al., 2007). There are currently no standardized methods for characterizing a food environment, and thus assessment and development of appropriate food environment measures is warranted (Caspi et al. 2012b; Lytle, 2009).

Perception measures based on surveys and self-report have been used increasingly to characterize food environments (Moore et al., 2008a, 2008b; Sharkey, 2009; McKinnon et al., 2009). Such measures have included an individual's perception of the availability of healthy food items in his/her neighborhood (Moore et al., 2008a, 2008b; Freedman and Bell, 2009; Zenk et al., 2009; Gustafson et al., 2011; Williams et al., 2011; Moore et al., 2012), as well as information on the individual's perceived presence of different retail food outlets (Zenk et al., 2009; Gustafson et al., 2011; Williams et al., 2011; Caspi et al., 2012b). In a recent review by Caspi and colleagues, studies using measures of perceived food environments have shown significant associations with dietary outcomes (Caspi et al. 2012a; Sharkey et al. 2010; Moore et al. 2009, 2008b; Inglis et al. 2008), whereas studies utilizing GIS-based measures of retail outlet presence or density have shown mixed and varying relationships (Caspi et al. 2012b).

Methodological decisions regarding how to define food environments, including geographic boundaries and contexts, could have a significant role in deciphering inconsistent findings among studies. Using GIS, food environments have typically been characterized by geographic "neighborhood" boundaries defined as census tracts, block groups and/or Euclidean or network buffers centered on some point of reference (e.g., population-weighted centroids or home addresses) (Charreire et al., 2010; Van Meter et al., 2011; Thornton et al., 2011). However, the use of such boundaries has notable limitations, including the modifiable areal unit problem, which can bias research findings based on the choice (i.e., number, size and shape) of boundaries for areal units (Openshaw, 1983; Christian, 2012). In addition, one cannot assume that all individuals conceptualize and/or interact within their environments similarly. For instance, in neighborhood perception studies utilizing mental maps, researchers found that an individual's perceived neighborhood can cover many different spaces and produce different boundaries based on age, race, class, gender and various other factors (Coulton et al., 2001, 2012). In a recent activity-based food environment study, researchers found that individuals encountered very different food environments in their daily travels compared to those located within or near their residential-defined neighborhood (Christian, 2012). Thus, how a person may operationalize and

perceive his/her neighborhood food environment could vary based on that person's daily routine and sociodemographic factors.

Despite challenges in defining and characterizing a food environment, only a handful of studies have examined differences in perceived and objective environments (Caspi et al., 2012a; Moore et al., 2008a, 2008b, 2012; Zenk et al., 2009; Gustafson et al., 2011; Williams et al., 2011; Giskes et al., 2007; Freedman and Bell, 2009). Of these studies, only two directly compared perceived and GIS-based presence (availability) of retail food outlets (Caspi et al., 2012a; Williams et al., 2011); the primary focus of these studies was the identification of traditional food outlets (i.e., supermarkets) within a pre-specified neighborhood boundary.

In this study, we sought to provide an in-depth comparison of GIS-based and perceived presence of retail food outlets in a sample of adults living in an eight-county region of South Carolina. In doing so, we aimed to determine what retail food outlets are available in an individual's neighborhood, as defined by field-validated GIS (as the gold standard), and to what extent individuals are aware of the presence of these food outlets via survey. Aims of our analyses included: 1) to examine whether the objective presence of retail food outlets within a standard 1-mile (~1.6 km) buffer used to define an individual's GIS-based neighborhood is accurately reflected in the perceived presence of retail food outlets within a 1-mile or a 20-minute walk from an individual's home; 2) to conduct sensitivity analyses by varying defined GIS-based neighborhoods, utilizing 2-, 3- and 5-mile buffers to examine changes in agreement (i.e., percent agreement and Kappa statistics) while keeping the perception buffer the same (1 mile or a 20-minute walk); and 3) to examine the accuracy between perceived and GIS-based presence by a sociodemographic factor, specifically urban or non-urban neighborhood designation.

Findings from this study could contribute to exploring whether measures of perceived availability of food retail outlets are viable alternative measures to GIS-based measures in food environment studies. In addition, this work could contribute to refining methods that researchers and policymakers use to describe a person's perception of his/her food environment and whether an individual's perceptions are adequate to detect changes in the retail food environment resulting from food access interventions, policy initiatives and associations with diet and weight outcomes.

2. Methods

This was a cross-sectional, non-experimental research study utilizing survey responses from 939 primary household food shoppers conducted in the spring of 2010, along with corresponding GIS-based measures of the respondents' food environments, within an eight-county region in South Carolina. This was a supplemental study related to a larger research effort focused on developing measures of the built nutritional environment (Liese et al., 2010, 2013a) and examining perceptions, shopping behaviors and diet among residents of the eight-county study region (Ma et al., 2013; Liese et al., 2013b). This study was approved by the University of South Carolina (USC) Institutional Review Board.

2.1 Study Region

The study area consisted of a contiguous geographical region encompassing eight counties (seven non-urban and one urban) in South Carolina (Fig. 1). The one urban county, Richland, contains the state capital, Columbia. The seven non-urban counties (Calhoun, Chester, Clarendon, Fairfield, Kershaw, Lancaster and Orangeburg) comprised the remaining study area.

2.2 Study Participants

Data collection occurred between April and July of 2010. Recruitment of study participants was geographically based and was designed to achieve spatial coverage of the entire study area. Specifically, selection was done through a random selection of landline telephone numbers with listed addresses restricted to 64 eligible ZIP codes within the study area, with a goal of 15 respondents per ZIP code. Recruitment calls were made by the interviewing staff at USC's Survey Research Laboratory. Respondents were screened for the eligibility criteria, including being 1) at least 18 years of age, 2) the primary food shopper of the household, 3) English speaking and 4) living in the eight-county study area. Of the 2,477 household telephone numbers screened, a total of 968 residents were eligible and completed the interview; there were 553 refusals, 377 ineligibles and 579 residents of non-contact, unknown or other status. Applying response rate formula 4 as outlined by the American Association for Public Opinion Research (American Association for Public Opinion Research (AAPOR), 2011), we estimated a response rate of 47%. The AAPOR is the leading professional organization of public opinion and survey research, providing standardized definitions and methodologies for the comparison of response rates for different surveys in the United States (AAPOR, 2011). Our survey response rate was very comparable to the 49% rate among landline households that was achieved in a recent evaluation of the Behavioral Risk Factor Surveillance System (BRFSS) landline response rates conducted in 18 US states (Hu et al., 2011). In our analyses, 11 participants were removed because of missing perception data, and 18 were removed because of errors in geocoding addresses. The final analytic sample included 939 participants.

2.3 Perception Measures

Perceived presence of a food retail outlet was obtained using a set of newly developed, validated questions (Table 1) (Ma et al., 2013). Specifically, participants were asked "which of the following stores, if any, are located in your neighborhood, defined as a 1-mile buffer or 20-minute walk around your home?" Stores listed included six types of retail food outlets, including supermarkets, supercenters, convenience stores, drug and pharmacy stores, dollar and variety stores and fast food restaurants. Response options were dichotomous: perceived presence was reported as yes or no. The test-retest reliability of the questions determined by Phi coefficients ranged from 0.58 for convenience stores to 0.96 for supermarkets (Ma et al., 2013). A person's neighborhood was defined as a 1-mile buffer or 20-minute walk around their home based on validated perception questions published previously (Moore et al., 2008a, 2008b; Mujahid et al., 2007). In analyses, supermarkets and supercenters (defined as any large retail store that sells both groceries and general merchandise, e.g., Wal-Mart or Target) were aggregated on the basis that supermarkets and supercenters typically represent

food outlets providing access to nutritious food in greater variety and affordability and higher quality (Franco et al., 2008; Block and Kouba, 2006). This classification was used previously by the US Centers for Disease Control and Prevention in their 2013 State Indicator Report on Fruits and Vegetables (Centers for Disease Control and Prevention, 2013).

2.4 GIS-Based Measures

The GIS-based presence of retail food outlets was determined using each respondent's home address as the point of reference, with varying (1-, 2-, 3- and 5-mile) Euclidean and street network buffers representing neighborhood boundaries (Fig. 2). The buffers were selected to provide a reasonable range of distances that could accommodate both sensitivity analyses and any urban vs. non-urban differences in the presence of retail food outlets. Overall, in our sample, the average distance to the nearest retail food outlet ranged from 3 to 8 miles. For urban participants, this average distance ranged from 1 to 2 miles, and for non-urban participants, it ranged from 3 to 9 miles. Thus, a 1-mile network buffer was selected as the standard comparison to perception measures, and incremental increases up to 5 miles were examined. For each outlet type, presence was determined by linking geospatial retail food environment data that were validated in 2009. Methods for creating the GIS-based retail food environment dataset have been published elsewhere (Liese et al., 2010, 2013a). Dichotomous variables representing presence (yes or no) for all food outlet types were created.

Residential addresses were geocoded using address-match geocoding in ArcGIS Network Analyst 10.0 (ESRI Redlands, CA 2010). Street network data were provided by StreetMap Premium within ArcGIS. Of the 968 addresses provided, 950 (98.1%) were matched to a corresponding street address on the street network in ArcGIS and assigned a location. In addition, a 15-meter side offset was selected to place addresses on the correct side of the road based on the methods of Cayo and Talbot (Cayo and Talbot, 2003). Eighteen (1.9%) participant addresses were unmatched. Of these, 5 were PO boxes and 13 had street names that did not exist in the street network data provided. The participants who provided these addresses were excluded from analyses.

2.5 Individual Characteristics

Demographic characteristics were determined based on questions from BRFSS (Centers for Disease Control and Prevention BRFSS, 2011) and included age, sex, race/ethnicity, education, employment status, utilization of the Supplemental Nutrition Assistance Program (SNAP), marital/partner status and number of individuals living in the home. Each survey respondent was also classified individually with respect to urbanicity—urban or non-urban—using the 2010 US Census—defined urban classification via a point-in-polygon operation within ArcGIS (United States Census Bureau, 2010). Specifically, a resident's home address was geocoded, and the US Census tract in which their home is located was determined. An urban area was determined as any census tract classified as either “an urbanized area of 50,000 or more people or an urban cluster consisting of at least 2,500 and less than 50,000 people.” All census tracts not designated as urban were considered to be non-urban.

2.6 Statistical Analyses

All statistical analyses were conducted using SAS 9.3 (SAS Institute, Inc., Cary, NC, 2012). First, the percent agreement was calculated to represent a basic measure of the proportion of respondents that accurately perceived the presence or absence of a food outlet type in their corresponding GIS-based neighborhood food environment when there was an actual food outlet presence or absence, respectively. Kappa statistics were then calculated to examine whether the objective GIS-based presence of retail food outlets was reflected by an individual's perceptions, after accounting for the effects of chance agreement alone (Cohen, 1960; Viera and Garrett, 2005).

For all statistics, GIS-based presence was considered the gold standard. 95% confidence intervals (CIs) were calculated for all statistics by approximating the binomial distribution with a normal distribution. For ease of interpretation of the Kappa statistic values, the following guidelines as outlined by Landis and Koch and other studies were used: 0.81–1.00, almost perfect; 0.61–0.80, substantial; 0.41–0.60, moderate; 0.21–0.40, fair; 0–0.20, slight; and <0, poor (Landis and Koch, 1977; Viera and Garrett, 2005).

When conducting sensitivity analyses, agreement statistics were constructed, varying by GIS-based neighborhood buffer sizes (i.e., 1-, 2-, 3-, and 5-mile buffers), while the measure of perceived presence remained at 1 mile (or a 20-minute walk). Differences between statistics by buffer sizes were then assessed using non-overlapping CIs. Thus, if the CIs for two statistics did not overlap, the values were considered significantly different. Finally, the agreement between perceived and GIS-based presence was compared for individuals living in urban and non-urban neighborhoods. Fisher's exact tests were performed to determine differences between groups.

3. Results

Descriptive statistics of the study sample are displayed in Table 2. Of the 939 participants, the majority were female (79%), non-Hispanic white (67%) and lived in non-urban neighborhoods (80%). The mean age for all individuals was ~58 years. Twelve percent of participants did not have a high school diploma, 22% were unemployed, 8% received SNAP benefits and 65% had a spouse or partner living in the household. On average, individuals lived with 2.5 household members. Ninety-four percent of all participants used their own personal vehicle for primary food shopping.

Thirty-two percent of individuals reported having a supermarket in their neighborhood, defined as a 1-mile buffer around their home, whereas only 19% had a GIS-verified supermarket within a 1-mile Euclidean-based buffer, and 12% had one within a 1-mile network-based buffer (Table 3). Similar discrepancies were observed for convenience stores (54% vs. 36% and 29%, respectively), drug and pharmacy stores (29% vs. 16% and 13%, respectively), dollar and variety stores (39% vs. 22% and 15%, respectively) and franchised fast food restaurants (27% vs. 22% and 15%, respectively). For all food outlet types, the Euclidean-based buffer was closer in percentage value to the perceived data as compared to the network-based buffer.

Participants who had a specific retail food outlet located within 1 mile of their home were aware of its presence or absence, as indicated by percent agreements ranging from 72% for convenience stores to 86% for fast food restaurants using a Euclidean-based 1-mile buffer (Table 4). Percent agreements ranged from 68% to 83% using a network-based 1-mile buffer. Increasing the buffer size did not result in any significant changes in percent agreement using either Euclidean- or network-based buffers. Thus, there seemed to be no advantage in comparing the perceived survey results to a larger GIS-based neighborhood buffer size. Kappa values ranged from 0.44 (moderate agreement) for convenience stores to 0.61 (substantial agreement) for fast food restaurants using a 1-mile Euclidean-based buffer. Kappa values were lower when using a 1-mile network-based buffer, ranging from 0.35 (fair agreement) for dollar and variety stores to 0.50 (moderate agreement) for fast food restaurants. In addition, the Kappa statistic for the presence of supermarkets was significantly lower when using a 1-mile network-based buffer (0.38, fair agreement) compared to a 1-mile Euclidean-based buffer (0.52, moderate agreement).

When using different GIS-based neighborhood buffer sizes (i.e., 2-, 3- and 5-mile) as objective measures of presence, there were some statistically significant differences between Kappa values compared to the standard 1-mile Euclidean- or network-based neighborhood buffer sizes (Table 4). Specifically, agreement for the presence of a supermarket, drug and pharmacy store or dollar and variety store was significantly higher when using either the 2- or 3-mile network-based buffer sizes compared to the 1-mile network-based buffer, indicating that the 1-mile buffer may underestimate a participant's perceived neighborhood. In addition, agreement for the presence of a fast food restaurant was significantly higher when using the 2-mile network-based buffer compared to the 1-mile buffer.

We also noted that the measures of agreement for the 3- and 5-mile buffers were significantly lower than those for the 1-mile Euclidean- and network-based buffers. However, this decrease in agreement is most likely due to the inevitable increase in the number of stores introduced mathematically into the denominator when calculating agreement statistics for much larger GIS-based buffer sizes. Thus, these findings are not necessarily comparable to the 1-mile buffer size after a certain mathematical threshold.

Finally, agreement was examined by stratifying participants by urban or non-urban designation (Table 5). Overall, the percent agreement was significantly lower for supermarkets and dollar and variety stores in urban participants compared to non-urban participants using either 1-mile Euclidean- or network-based buffers. In addition, the percent agreement was significantly lower for drug and pharmacy stores and fast food restaurants in urban participants compared to non-urban participants using a 1-mile network-based buffer. However, there were few significant differences between urban and non-urban participants when examining Kappa values, as only the agreement for drug and pharmacy stores differed significantly using either Euclidean- or network-based buffers, indicating a significantly lower agreement in non-urban participants compared to urban participants. Lastly, urban participants had a significantly lower chance-adjusted agreement for dollar and variety stores compared to non-urban participants using the network-based buffer.

4. Discussion

In this study, we examined the relationship between perceived and GIS-based presence of retail food outlets in a sample of adults living in South Carolina. Overall, an individual's perception of the presence or absence of a food store located within 1 mile of his/her home was moderately agreeable with objective, GIS-based measures of food outlet presence. Kappa values ranged from 0.44 to 0.61 when comparing a participant's perceived presence of food outlets to the GIS-based presence of food outlets using a 1-mile Euclidean-based neighborhood buffer and 0.35 to 0.50 using a 1-mile network-based buffer.

Our findings also suggest that the performance of an individual's perceptions was most optimal when using a 2- or 3-mile geographic-based neighborhood boundary. Kappa values for comparisons of these perceptions to standard 1-mile Euclidean- or network-based neighborhood buffer sizes were significantly higher for supermarkets, drug and pharmacy stores and dollar and variety stores when using either 2- or 3-mile network-based buffer sizes compared to a 1-mile network-based buffer. In addition, the agreement for fast food restaurants was significantly higher using the 2-mile network-based buffer. These findings suggest that researchers should consider the neighborhood boundaries selected as a reference in food environment measures. Although Euclidean-based measures were not significantly different, when analyzing network-based measures, researchers should consider the most appropriate boundary size: it could be that a 1-mile network-based buffer does not necessarily best characterize a participant's perceived neighborhood.

To the best of our knowledge, this is the first study to explore the impact of varying neighborhood buffer sizes on the relationship between perceived and actual presence of various types of food outlets. However, we do not wish to overstate the contribution of our findings, as we did not have varied neighborhood boundaries when administering the perception questionnaire. A future study comparing both perceived and GIS-based measures and utilizing varied neighborhood boundaries would be ideal.

Agreement statistics also differed significantly when considering urbanicity. Specifically, the percent agreement was significantly lower for supermarkets and dollar and variety stores in urban participants compared to non-urban participants using either 1-mile Euclidean- or network-based buffers. In addition, the percent agreement was significantly lower for drug and pharmacy stores and fast food restaurants in urban participants using a 1-mile network-based buffer. However, given that we observed few significant differences by urbanicity after accounting for chance agreement via Kappa statistics, this finding suggests that for supermarkets, convenience stores and franchised fast food outlets, residents of urban and non-urban areas do not differ substantially in terms of their reported agreement.

Two recent studies compared perceived and GIS-based presence of food outlets (Williams et al., 2011; Caspi et al., 2012b). In a sample of women aged 18–65 years in Melbourne, Australia, Williams et al. found that the match (percent agreement) between perceived and objective food environments was poor, reporting that approximately 50% of women had accurate reports between their perceptions and objective measures of supermarket presence within 800 m (~0.5 miles) of their home (Williams et al., 2011). For fast food stores, this

agreement was only 40%. This outcome is much different than that in our study, where we report percent agreements of 78% using a 1-mile network-based buffer size and 82% using a 1-mile Euclidean-based buffer. Possible discrepancies between our results and those of Williams et al. may be traced to the nature of the perception questions used and the choice of GIS-based measure. In our study, we asked study participants to think of their neighborhood environment specifically as a 1-mile buffer or 20-minute walk around their home, whereas participants in the study by Williams et al. were asked “are the following within walking distance of your home?” without any guide to what is meant by walking distance. Moreover, Williams et al. classified participants via GIS as having or not having each type of store nearby using an 800-m (~0.5-mile) definition for walkable distance.

In another study, Caspi et al. reported that 69% of residents matched, via percent agreement, between objective and perceived presence of a supermarket within 1 kilometer (~0.6 miles) in a sample of low-income housing residents in three urban areas in the greater Boston area. This may suggest that Caspi et al. used a buffer size that was too small to optimize concordance between a person’s perceived and objectively measured food environment. Moreover, Caspi et al. chose their cut point for a neighborhood buffer at 1 kilometer because the researchers were concerned about artificially high levels of discordance based on previous buffers used in the literature and because most of their participants reported a supermarket within walking distance (Caspi, Kawachi, Subramanian, Adamkiewicz, & Sorensen 2012a). It is important to point out, however, that in our study, the majority of food shoppers traveled by car (>90%) and did not walk to food outlets, even in urban neighborhoods.

Our study contributes to food environment research by not only exploring the agreement between an individual’s perceived and actual presence of supermarkets and other food retail outlets but also examining how this relationship changes using different boundaries to define a person’s neighborhood. It could be that the Williams et al. and Caspi et al. studies using smaller cut points to define a person’s neighborhood may have affected the relationship between perception and the GIS-based environment. In our study, we found slight improvement, via significant increases in Kappa statistics, when increasing the network-based neighborhood definition from the 1-mile to the 2- or 3-mile buffer sizes. It is possible that individuals are not able to mentally conceptualize the extent of a 1-mile buffer around their home based on personal and behavioral factors. Thus, additional studies are needed that vary the spatial frame of the perception questions (i.e., 2 miles, 3 miles, etc.) and subsequently compare respondents’ perceptions with validated GIS-based measures with the same buffers as those used in the perception questions.

Our study has several limitations. First, women constituted the majority of the sample because we selected the primary food shopper of the household. This may limit the generalizability of our findings. Second, our landline-based telephone sample yielded an age distribution with an average age in the middle-to-older category, which does not represent all residents. Third, the perception data were collected nearly 1 year after the completion of the validated field census. However, it should be noted that this time gap between data collection is comparable to those reported in other similar studies (Moore et al., 2008a; Gustafson et al., 2011).

Strengths of the study included the use of a set of reliable food environment questions examining the perceived presence of various types of food outlets (Ma et al., 2013). Reliability ranged from 0.58 to 0.96 in a test-retest analysis (Ma et al., 2013). Secondly, our measure of GIS-based presence was based on a validated field census of our study region (Liese et al., 2010, 2013a). In addition, our study area contained both urban and non-urban communities. Thus, these findings may be beneficial and comparable to new studies examining populations in the southeastern United States, where there is a mix of urban and non-urban neighborhoods. Williams et al. and Caspi et al. both examined residents living in urban communities only.

5. Conclusions

GIS is an important and useful tool for describing a food environment and relating this environment to an individual's diet, weight status and neighborhood characteristics; however, measures based on GIS are susceptible to a number of errors and inaccuracies (Liese et al., 2010, 2013a; Powell et al., 2011; Forsyth et al., 2010). This susceptibility is due partially to the fact that few datasets and resources have been conceived with the sole purpose of capturing food outlet access. Moreover, the effort to validate commercial and secondary datasets is often not feasible because of a lack of resources and the expense of research staff traveling into the field (Forsyth et al., 2010). There is debate as to whether it may be more economical and accurate to utilize perceptions measures, either alone or in tandem with GIS-based measurements (Moore et al., 2008a; Caspi et al., 2012a). For studies that are still in the planning stage, our results could suggest that a simple survey question alone may be advantageous, thereby saving a study team the effort of lengthy data cleaning and merging, ground truthing and/or GIS analyses. Moreover, our study demonstrates there is a moderate correspondence between what retail food outlets an individual perceives in his/her neighborhood compared to what is actually present, especially in terms of supermarkets. However, our study also points out that there is still room to evaluate appropriate neighborhood boundaries for both GIS-based and perceived presence of retail food outlets. More evidence is needed to determine whether to rely on perception-based measures of retail food outlets exclusively.

Acknowledgments

This project was supported by grant R21CA132133-02S1 from the National Cancer Institute. The contents of this article are solely the responsibility of the authors and do not necessarily represent the official views of the National Cancer Institute or the National Institutes of Health. The authors would like to thank James Hibbert for all GIS data management and mapping.

References

- American Association for Public Opinion Research (AAPOR). Standard definitions, final dispositions of case codes and outcome rates for surveys. American Association for Public Opinion Research; 2011.
- Block D, Kouba J. A comparison of the availability and affordability of a market basket in two communities in the Chicago area. *Public Health Nutrition*. 2006; 9:837–845. [PubMed: 17010248]
- Bodor JN, Rose D, Farley TA, Swalm C, Scott SK. Neighbourhood fruit and vegetable availability and consumption: the role of small food stores in an urban environment. *Public Health Nutrition*. 2008; 11:413–420. [PubMed: 17617930]

- Caspi CE, Kawachi I, Subramanian SV, Adamkiewicz G, Sorensen G. The relationship between diet and perceived and objective access to supermarkets among low-income housing residents. *Social Science & Medicine*. 2012a; 75:1254–1262. [PubMed: 22727742]
- Caspi CE, Sorensen G, Subramanian SV, Kawachi I. The local food environment and diet: a systematic review. *Health & Place*. 2012b; 18:1172–1187. [PubMed: 22717379]
- Cayo MR, Talbot TO. Positional error in automated geocoding of residential addresses. *International Journal of Health Geographics*. 2003; 2:10. [PubMed: 14687425]
- Centers for Disease Control and Prevention (CDC). [Accessed February 10, 2014] State indicator report on fruits and vegetables, 2013. 2013. Available from: <http://www.cdc.gov/nutrition/downloads/State-Indicator-Report-Fruits-Vegetables-2013.pdf>
- Centers for Disease Control and Prevention (CDC). Behavioral Risk Factor Surveillance System Survey Data. U.S. Department of Health and Human Services, Centers for Disease Control and Prevention; Atlanta, Georgia: 2011.
- Charreire H, Casey R, Salze P, Simon C, Chaix B, Banos A, Badariotti D, Weber C, Oppert JM. Measuring the food environment using geographical information systems: a methodological review. *Public Health Nutrition*. 2010; 13:1773–1785. [PubMed: 20409354]
- Christain J. Using geospatial technologies to explore activity-based retail food environments. *Spatial and Spatio-temporal Epidemiology*. 2012; 3:287–295. [PubMed: 23149325]
- Cohen J. A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*. 1960; 20:37–46.
- Coulton CJ, Jennings MZ, Chan T. How big is my neighborhood? Individual and contextual effects on perceptions of neighborhood scale. *American Journal of Community Psychology*. 2012; 51:140–150. [PubMed: 22886284]
- Coulton CJ, Korbin J, Chan T, Su M. Mapping residents' perceptions of neighborhood boundaries: a methodological note. *American Journal of Community Psychology*. 2001; 29:371–383. [PubMed: 11446289]
- Forsyth A, Lytle L, Riper DV. Finding food: Issues and challenges in using Geographic Information Systems to measure food access. *Journal of Transport and Land Use*. 2010; 3:43–65. [PubMed: 21837264]
- Franco M, Diez Roux AV, Glass TA, Caballero B, Brancati FL. Neighborhood characteristics and availability of healthy foods in Baltimore. *American Journal of Preventive Medicine*. 2008; 35:561–567. [PubMed: 18842389]
- Freedman DA, Bell BA. Access to healthful foods among an urban food insecure population: perceptions versus reality. *Journal of Urban Health*. 2009; 86:825–838. [PubMed: 19898939]
- Giskes K, Van Lenthe FJ, Brug J, Mackenbach JP, Turrell G. Socioeconomic inequalities in food purchasing: the contribution of respondent-perceived and actually (objectively measured) price and availability of foods. *Preventive Medicine*. 2007; 45:41–48. [PubMed: 17532036]
- Gustafson AA, Sharkey J, Samuel-Hodge CD, Jones-Smith J, Folds MC, Cai J, Ammerman AS. Perceived and objective measures of the food store environment and the association with weight and diet among low-income women in North Carolina. *Public Health Nutrition*. 2011; 14:1032–1038. [PubMed: 21324229]
- Hu SS, Balluz L, Battaglia MP, Frankel MR. Improving public health surveillance using a dual-frame survey of landline and cell phone numbers. *American Journal of Epidemiology*. 2011; 173:703–711. [PubMed: 21343246]
- Inglis V, Ball K, Crawford D. Socioeconomic variations in women's diets: what is the role perceptions of the local food environment? *Journal of Epidemiology and Community Health*. 2008; 62 (3): 191–197. [PubMed: 18272732]
- Jago R, Baranowski T, Baranowski JC, Cullen KW, Thompson D. Distance to food stores & adolescent male fruit and vegetable consumption: mediation effects. *International Journal of Behavioral Nutrition & Physical Activity*. 2007; 4:35. [PubMed: 17850673]
- Janse AJ, Gemke RJ, Uiterwaal CS, van dT I, Kimpfen JL, Sinnema G. Quality of life: patients and doctors don't always agree: a meta-analysis. *Journal of Clinical Epidemiology*. 2004; 57:653–661. [PubMed: 15358393]

- Jeffery RW, Baxter J, McGuire M, Linde J. Are fast food restaurants an environmental risk factor for obesity? *International Journal of Behavioral Nutrition & Physical Activity*. 2006; 3:2. [PubMed: 16436207]
- Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics*. 1977; 33:159–174. [PubMed: 843571]
- Laraia BA, Siega-Riz AM, Kaufman JS, Jones SJ. Proximity of supermarkets is positively associated with diet quality index for pregnancy. *Preventive Medicine*. 2004; 39:869–875. [PubMed: 15475018]
- Larson NI, Story MT, Nelson MC. Neighborhood environments: disparities in access to healthy foods in the U.S. *American Journal of Preventive Medicine*. 2009; 36:74–81. [PubMed: 18977112]
- Liese AD, Colabianchi N, Lamichhane AP, Barnes TL, Hibbert JD, Porter DE, Nichols MD, Lawson AB. Validation of 3 food outlet databases: completeness and geospatial accuracy in rural and urban food environments *American Journal of Epidemiology*. 2010; 172:1324–1333.
- Liese AD, Barnes TL, Lamichhane AP, Hibbert JD, Colabianchi N, Lawson AB. Characterizing the food retail environment: Impact of count, type and geospatial error in two secondary data sources. *Journal of Nutrition Education and Behavior*. 2013a; 45:435–442. [PubMed: 23582231]
- Liese AD, Bell BA, Barnes TL, Colabianchi N, Hibbert JD, Blake CE, Freedman DA. Environmental influences on fruit and vegetable intake: Results from a path analytic model. *Public Health Nutrition* Nov. 2013b; 6:1–10.
- Lytle L. Measuring the food environment: state of the science. *American Journal of Preventive Medicine*. 2009; 36 (4):S124–S133. [PubMed: 19285203]
- Ma X, Barnes TL, Freedman DA, Bell BA, Colabianchi N, Liese AD. Test-retest reliability of a questionnaire measuring perceptions of neighborhood food environment. *Health & Place*. 2013; 21:65–69. [PubMed: 23434497]
- McKinnon RA, Reedy J, Morrisette MA, Lytle LA, Yaroch AL. Measures of the food environment: a compilation of the literature, 1990–2007. *American Journal of Preventive Medicine*. 2009; 36:S124–S133. [PubMed: 19285203]
- Moore LV, Diez Roux AV, Brines S. Comparing perception-based and geographic information system (GIS)-based characterizations of the local food environment. *Journal of Urban Health*. 2008a; 85:206–216. [PubMed: 18247121]
- Moore LV, Diez Roux AV, Nettleton JA, Jacobs DR Jr. Associations of the local food environment with diet quality--a comparison of assessments based on surveys and geographic information systems: the multi-ethnic study of atherosclerosis. *American Journal of Epidemiology*. 2008b; 167:917–924. [PubMed: 18304960]
- Moore LV, Diez Roux AV, Nettleton JA, Jacobs DR, Franco M. Fast-food consumption, diet quality, and neighborhood exposure to fast food: the multi-ethnic study of atherosclerosis. *American Journal of Epidemiology*. 2009; 170 (1):29–36. [PubMed: 19429879]
- Moore LV, Diez Roux AV, Franco M. Measuring availability of healthy foods: Agreement between directly measured and self-reported data. *American Journal of Epidemiology*. 2012; 175:1037–1044. [PubMed: 22273535]
- Mujahid MS, Diez Roux AV, Morenoff JD, Raghunathan T. Assessing the measurement properties of neighborhood scales: from psychometrics to econometrics. *American Journal of Epidemiology*. 2007; 165:858–867. [PubMed: 17329713]
- Openshaw, S. The Modifiable Areal Unit Problem. Vol. 38. Norwich, England: Geobooks; 1983.
- Paquet C, Daniel M, Kestens Y, Leger K, Gauvin L. Field validation of listings of food stores and commercial physical activity establishments from secondary data. *International Journal of Behavioral Nutrition & Physical Activity*. 2008; 5:58. [PubMed: 19000319]
- Powell LM, Han E, Zenk SN, Khan T, Quinn CM, Gibbs KP, Pugach O, Barker DC, Resnick EA, Myllyluoma J, Chaloupka FJ. Field validation of secondary commercial data sources on the retail food outlet environment in the U.S. *Health & Place*. 2011; 17:1122–1131. [PubMed: 21741875]
- Sharkey JR. Measuring potential access to food stores and food-service places in rural areas in the U.S. *American Journal of Preventive Medicine*. 2009; 36:S151–S155. [PubMed: 19285206]

- Thornton LE, Pearce JR, Kavanagh AM. Using geographic information systems (GIS) to assess the role of the built environment in influencing obesity: a glossary. *International Journal of Behavioral Nutrition & Physical Activity*. 2011; 8:71. [PubMed: 21722367]
- U.S. Department of Agriculture (USDA). United States Department of Agriculture Food Desert Definitions. U.S. Department of Agriculture, Economic Research Service; 2014. Available from: <http://apps.ams.usda.gov/fooddeserts/foodDeserts.aspx> [Accessed February 10, 2014]
- United States Census Bureau. Urban Rural Classification. [Accessed on February 10, 2014] United States Census Bureau. 2010. Available from: <http://www.census.gov/geo/reference/ua/urban-rural-2010.html>
- Van Meter ME, Lawson AB, Colabianchi N, Nichols M, Hibbert J, Porter D, Liese AD. Spatial accessibility and availability measures and statistical properties in the food environment. *Spatial and Spatio-temporal Epidemiology*. 2011; 2:35–47. [PubMed: 21499528]
- Viera AJ, Garrett JM. Understanding interobserver agreement: The Kappa statistic. *Family Medicine*. 2005; 37:360–363. [PubMed: 15883903]
- Wang MC, Kim S, Gonzalez AA, MacLeod KE, Winkleby MA. Socioeconomic and food-related physical characteristics of the neighbourhood environment are associated with body mass index. *Journal of Epidemiology & Community*. 2007; 61:491–498.
- Williams LK, Thornton L, Ball K, Crawford D. Is the objective food environment associated with perceptions of the food environment? *Public Health Nutrition*. 2011; 15:291–298. [PubMed: 21835079]
- Zenk SN, Schulz AJ, Lachance LL, Mentz G, Kannan S, Ridella W, Galea S. Multilevel correlates of satisfaction with neighborhood availability of fresh fruits and vegetables. *Annals of Behavioral Medicine*. 2009; 38:48–59. [PubMed: 19809859]

Highlights

- Individuals moderately perceived the presence of food outlets in their neighborhood
- Agreement statistics were affected by neighborhood boundaries and urbanicity
- Neighborhood boundaries should be considered in future food environment measures

Fig 1.
Study region.

Fig 2.
Example of 1-mile Euclidean and network buffers.

Table 1

Survey questions on perceptions of the food environment.

Which of the following stores, if any, are located in your neighborhood, defined as a 1-mile buffer or 20-minute walk around your home? ^a	
1	A supercenter, such as Wal-Mart or Target
2	A supermarket, such as Food Lion, Kroger, Publix or Piggly Wiggly
3	A convenience store with or without a gas station attached
4	A freestanding drugstore or pharmacy store, such as CVS, Rite-Aid, Eckerd's or Walgreen's
5	A Dollar Variety, Dollar General, Dollar Store or Dollar Tree
6	A franchised fast food restaurant, including places like McDonald's, Subway or Taco Bell

^aResponse options are either "yes" or "no"

Table 2

Study participant demographics (N=939).

	%	Missing Data n
Age in years, mean (SD)	57.5 (14.6)	65
Age in years, %		65
<45	25	
45–64	44	
65	31	
Gender, %		0
Male	21	
Female	79	
Race/Ethnicity, %		67
Minority (non-Hispanic black, Hispanic or other)	33	
Non-Hispanic white	67	
Education, %		63
<High school diploma	12	
High school diploma	35	
Some college or higher	53	
Employment, %		62
Not employed	22	
Retired	35	
Employed	43	
Income per year, %		202
<\$20,000	29	
\$20,000–39,999	27	
\$40,000–59,999	18	
\$60,000	26	
Supplement Nutrition Assistance Program receipt, %		63
Yes	8	
No	92	
Spouse or partner in household, %		58
Yes	65	
No	35	
# of household members, mean (SD)	2.5 (1.4)	68
Transportation for primary shopping, %		13
Personal car/vehicle	94	
Ride with family/friends	5	
Public transportation	0.3	
Walk, bike or taxi	0.7	
Urbanicity, %		0
Urban	20	
Non-urban	80	

Table 3

Perceived and GIS-based presence of food retail outlets using standard 1-mile neighborhood buffer sizes (N=939).

	Perception-based Yes, %	Euclidean-based 1-mile buffer Yes, %	Network-based 1-mile buffer Yes, %
Supermarkets	32	19	12
Convenience	54	36	29
Drug and pharmacy	29	16	13
Dollar and variety	39	22	15
Franchised fast food	27	22	15

Table 4

Agreement between perceived and GIS-based presence of food retail outlets by varying neighborhood buffer sizes (N=939).

	Percent agreement (95% CI)		Kappa statistic (95% CI)	
	Euclidean-based	Network-based	Euclidean-based	Network-based
Supermarkets				
1-mile buffer	82 (77 – 86)	78 (73 – 83)	0.52 (0.46 – 0.58)	0.38 (0.32 – 0.44) [†]
2-mile buffer	84 (79 – 88)	84 (80 – 88)	0.63 (0.57 – 0.68)	0.60 (0.55 – 0.66)*
3-mile buffer	79 (74 – 83)	83 (79 – 87)	0.54 (0.49 – 0.60)	0.61 (0.56 – 0.67)*
5-mile buffer	70 (65 – 75)*	75 (70 – 80)	0.43 (0.38 – 0.47)	0.49 (0.44 – 0.54)
Convenience				
1-mile buffer	72 (68 – 76)	68 (64 – 72)	0.44 (0.39 – 0.50)	0.38 (0.33 – 0.43)
2-mile buffer	71 (67 – 75)	72 (68 – 76)	0.41 (0.35 – 0.47)	0.45 (0.39 – 0.50)
3-mile buffer	65 (61 – 69)	70 (66 – 74)	0.27 (0.22 – 0.33)*	0.38 (0.32 – 0.44)
5-mile buffer	56 (51 – 60)*	59 (55 – 63)*	0.05 (0.02 – 0.09)*	0.14 (0.09 – 0.19)*
Drug and pharmacy				
1-mile buffer	80 (76 – 85)	79 (75 – 84)	0.46 (0.39 – 0.52)	0.41 (0.35 – 0.47)
2-mile buffer	82 (77 – 86)	82 (78 – 87)	0.55 (0.49 – 0.61)	0.54 (0.48 – 0.60)*
3-mile buffer	80 (75 – 84)	81 (77 – 86)	0.53 (0.47 – 0.59)	0.55 (0.49 – 0.60)*
5-mile buffer	71 (66 – 77)	74 (69 – 80)	0.40 (0.34 – 0.46)	0.48 (0.39 – 0.51)
Dollar and variety				
1-mile buffer	77 (73 – 82)	72 (68 – 77)	0.49 (0.43 – 0.54)	0.35 (0.30 – 0.41) [†]
2-mile buffer	80 (76 – 84)	80 (75 – 84)	0.58 (0.52 – 0.63)	0.55 (0.50 – 0.61)*
3-mile buffer	77 (73 – 82)	80 (76 – 84)	0.54 (0.48 – 0.59)	0.58 (0.53 – 0.64)*
5-mile buffer	66 (61 – 71)*	71 (66 – 76)	0.37 (0.32 – 0.42)*	0.43 (0.38 – 0.49)
Franchised fast food				
1-mile buffer	86 (81 – 90)	83 (78 – 88)	0.61 (0.55 – 0.67)	0.50 (0.44 – 0.57)
2-mile buffer	84 (79 – 89)	86 (81 – 90)	0.61 (0.56 – 0.67)	0.64 (0.58 – 0.69)*
3-mile buffer	79 (74 – 84)	83 (78 – 88)	0.54 (0.48 – 0.59)	0.60 (0.55 – 0.66)

Percent agreement (95% CI)		Kappa statistic (95% CI)	
Euclidean-based	Network-based	Euclidean-based	Network-based
5-mile buffer	66 (60 – 72) *	72 (67 – 78) *	0.44 (0.39 – 0.49)

* Non-overlapping confidence intervals were used to indicate that the statistics were significantly different vs. the 1-mile neighborhood buffer (reference group)

† Non-overlapping confidence intervals were used to indicate that the statistics were significantly different vs. the Euclidean-based buffer at the corresponding neighborhood buffer size

Agreement between perceived and GIS-based presence of food retail outlets by urbanicity using the standard 1-mile neighborhood buffer sizes (N=939).

Table 5

	Percent agreement (95% CI)				Kappa statistic (95% CI)			
	Euclidean-based		Network-based		Euclidean-based		Network-based	
	Urban	Non-urban	Urban	Non-urban	Urban	Non-urban	Urban	Non-urban
Supermarkets	71 (63 – 78)	84 (78 – 90) *	56 (48 – 65)	83 (77 – 89) *	0.40 (0.28 – 0.52)	0.42 (0.33 – 0.50)	0.25 (0.16 – 0.33)	0.33 (0.25 – 0.41)
Convenience	76 (69 – 83)	71 (66 – 75)	70 (62 – 77)	68 (63 – 73)	0.37 (0.22 – 0.52)	0.40 (0.34 – 0.46)	0.34 (0.21 – 0.46)	0.34 (0.28 – 0.40)
Drug and pharmacy	72 (65 – 80)	82 (76 – 89)	68 (60 – 76)	83 (76 – 89) *	0.44 (0.33 – 0.56)	0.25 (0.16 – 0.33) *	0.39 (0.29 – 0.50)	0.21 (0.13 – 0.30) *
Dollar and variety	66 (57 – 74)	80 (75 – 86) *	53 (44 – 62)	78 (72 – 83) *	0.32 (0.20 – 0.45)	0.49 (0.42 – 0.55)	0.16 (0.07 – 0.26)	0.39 (0.33 – 0.46) *
Franchised fast food	78 (70 – 85)	88 (82 – 94)	67 (59 – 76)	87 (81 – 93) *	0.50 (0.37 – 0.63)	0.48 (0.39 – 0.57)	0.37 (0.26 – 0.49)	0.41 (0.32 – 0.51)

* Significant difference between groups (urban vs. non-urban), Fisher exact test, two-tailed, $p < 0.05$